

# Deep Koopman Analysis

Estimer les éléments invariants de l'opérateur de Koopman pour linéariser les systèmes dynamiques

**Joris, Sylviane, Vicky**

Présenté le 04/03/2024

## Résumé

L'objectif de ce rapport est de présenter de manière concise un résumé des travaux [1] et [2] dans l'analyse de l'opérateur de Koopman. L'objectif final est d'implémenter un réseau neuronal profond afin d'estimer l'opérateur de Koopman dans des cas simples.

# Table des matières

<b>1</b>	<b>Opérateur de Koopman</b>	<b>2</b>
1.1	Systèmes dynamiques . . . . .	2
1.2	Définition . . . . .	3
1.3	Espace invariant de Koopman . . . . .	3
<b>2</b>	<b>Architecture</b>	<b>4</b>
2.1	Fonctions de perte . . . . .	5
<b>3</b>	<b>Résultats et applications</b>	<b>6</b>
3.1	Résultats théoriques . . . . .	6
3.1.1	Systèmes dynamiques pilotés par les données . . . . .	6
3.1.2	Théorie des opérateurs de Koopman . . . . .	6
3.1.3	Koopman pour les systèmes à spectre continu . . . . .	7
3.1.4	Identification les fonctions propres de Koopman . . . . .	7
3.2	Résultats appliqués à des exemples . . . . .	8
3.2.1	Modèle simple avec spectre discret . . . . .	8
3.2.2	Pendule non linéaire avec spectre continu . . . . .	8
3.2.3	Écoulement de fluide non linéaire à haute dimension . . . . .	8
3.2.4	Comparaison des exemples . . . . .	8
3.3	Applications . . . . .	9

# 1 Opérateur de Koopman

## 1.1 Systèmes dynamiques

Que ce soit en physique, en biologie ou dans d'autres domaines, un *Système dynamique* est simplement un système qui évolue dans le temps et est régi par un ensemble d'états  $x \in \mathbb{R}^p$  satisfaisant une équation différentielle du premier ordre :

$$\dot{x} = f(x, t)$$

où la notation  $\dot{x}$  désigne la dérivée temporelle du vecteur  $x$  à un instant donné  $t$ .

L'étude de tels systèmes s'avère difficile dans deux cas :

- On étudie un système d'un point de vue continu
- La relation n'est pas toujours linéarisable

Pour simplifier le premier obstacle et pour simplifier l'analyse du système, on peut discrétiser le système en proposant l'approximation suivante :

$$x_{t+1} \approx x_t + \dot{x} := \tilde{f}(x_t)$$

Afin de résoudre le problème de non-linéarité de la fonction  $\tilde{f}$ , Koopman [3] propose en 1931 de changer de paradigme.

**Exemple 1.1** *On s'intéresse au système défini par :*

$$\begin{cases} \dot{x}_1 = \mu x_1 \\ \dot{x}_2 = \lambda(x_2 - x_1^2) \end{cases} \quad (1)$$

En posant que  $y_1 = x_1$ ,  $y_2 = x_2$  et  $y_3 = x_1^2$ , il est possible de linéariser le système en dimension finie

$$\frac{d}{dt} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \mu & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & 2\mu \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

**Exemple 1.2** *Ici, on a un exemple, où il n'est pas possible de linéariser le système en dimension finie :*

$$\dot{x} = x^2$$

On peut le linéariser en dimension infinie en posant que

$$y_1 = x, y_2 = x^2, y_3 = x^3, \dots$$

Donc, on a :

$$\frac{d}{dt} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \vdots \end{pmatrix}$$

## 1.2 Définition

Plutôt que d'étudier directement l'espace d'états  $x \in \mathbb{R}^p$ , on propose d'étudier l'espace de Hilbert des mesures de l'état du système  $y = g(x)$ . On peut penser à  $g(x) := x_i$  pour l'étude d'une dimension du système ou encore  $g(x) := \frac{1}{2}m\dot{x}$  mesure de l'énergie cinétique du système.

On définit alors l'*Opérateur de Koopman* associé à un système dynamique par :

$$\mathcal{K}g := g \circ \tilde{f} \tag{2}$$

On obtient alors :

$$\mathcal{K}g(x_t) = g \circ \tilde{f}(x_t) \tag{3}$$

$$= g(\tilde{f}(x_t)) \tag{4}$$

$$= g(x_{t+1}) \tag{5}$$

**Propriété :** L'opérateur de Koopman est linéaire

L'opérateur de Koopman propose alors un changement de paradigme : On passe d'un système d'états de dimension finie non linéaire à un opérateur linéaire dans un espace de dimensions infini.

## 1.3 Espace invariant de Koopman

L'opérateur de Koopman étant un opérateur d'un espace de Hilbert, on peut alors s'intéresser aux valeurs et vecteurs (fonctions) propres de cet opérateur.

**Exemple 1.3** Pour le système de l'équation (1), on s'intéresse à la fonction  $g(x_1, x_2) \mapsto x_1$ . On montre alors que :

$$\mathcal{K}g(x_t) = g(x_{t+1}) = (1 + \mu)g(x_t) \tag{6}$$

Donc  $g$  est fonction propre de  $\mathcal{K}$  associé à la valeur propre  $(1 + \mu)$ .

On s'intéresse alors à la recherche alors d'espaces de fonctions  $y_1, \dots, y_n$  fonctions propres de l'opérateur de Koopman (ainsi que les valeurs propres associées).

Afin de pouvoir mesurer précisément l'état du système, on peut imposer que  $y_1 = x_1, \dots, y_p = x_p$  afin de pouvoir conserver la description de l'état du système à l'instant  $t$ .

**Propriété :** Par linéarité de l'opérateur, si  $y_1, \dots, y_n$  sont des fonctions propres de l'opérateur de Koopman, alors  $\text{span} \{y_1, \dots, y_n\}$  est un espace invariant de Koopman.

**Exemple 1.4** Dans l'exemple précédent, on pose :

$$\begin{cases} y_1 = x_1 \\ y_2 = x_2 \\ y_3 = x_1^2 \end{cases} \quad (7)$$

Dans ce cas,  $\text{span} \{y_1, y_2, y_3\}$  est un espace invariant de Koopman.

## 2 Architecture

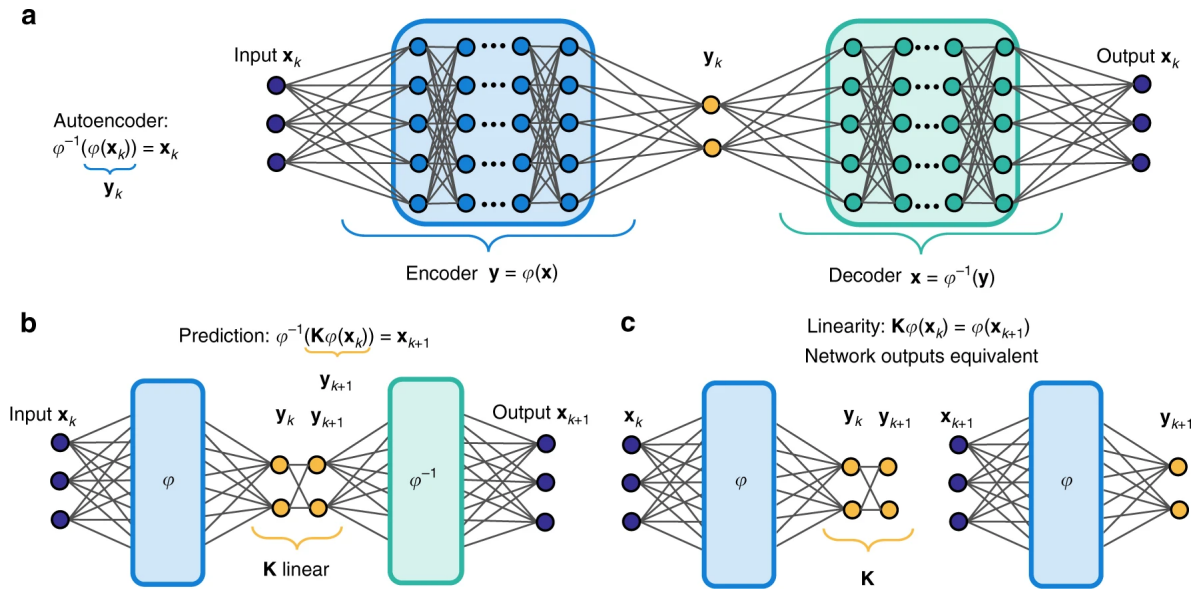


FIGURE 1 – L'architecture du modèle

Voici l'architecture de Deep Koopman pour identifier les fonctions propres de Koopman  $\varphi(x)$ . Le réseau est basé sur un auto-encodeur qui est capable d'identifier les coordonnées intrinsèques  $y = \varphi(x)$  et de décoder ces coordonnées pour récupérer  $x = \varphi^{-1}(y)$ .

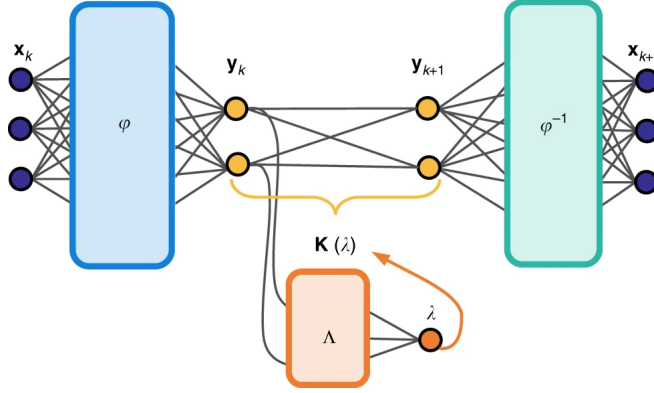


FIGURE 2 – réseau auxiliaire

L'ajout d'un réseau auxiliaire sert à identifier le spectre des valeurs propres continues  $\lambda$ . Cela facilite une réduction de dimension agressive dans l'auto-encodeur, évitant ainsi le besoin de plus hautes harmoniques de la fréquence qui sont générées par la non-linéarité.

## 2.1 Fonctions de perte

Voici les fonctions de perte utilisées :

$$\mathcal{L} = \alpha_1(\mathcal{L}_{\text{recon}} + \mathcal{L}_{\text{pred}}) + \mathcal{L}_{\text{lin}} + \alpha_2 \mathcal{L}_{\infty} + \alpha_3 \|\mathbf{W}\|_2^2 \quad (8)$$

$$\mathcal{L}_{\text{recon}} = \|\mathbf{x}_1 - \varphi^{-1}(\varphi(\mathbf{x}_1))\|_{\text{MSE}} \quad (9)$$

$$\mathcal{L}_{\text{pred}} = \frac{1}{S_p} \sum_{m=1}^{S_p} \|\mathbf{x}_{m+1} - \varphi^{-1}(K^m \varphi(\mathbf{x}_1))\|_{\text{MSE}} \quad (10)$$

$$\mathcal{L}_{\text{lin}} = \frac{1}{T-1} \sum_{m=1}^{T-1} \|\varphi(\mathbf{x}_{m+1}) - K^m \varphi(\mathbf{x}_1)\|_{\text{MSE}} \quad (11)$$

$$\mathcal{L}_{\infty} = \|\mathbf{x}_1 - \varphi^{-1}(\varphi(\mathbf{x}_1))\|_{\infty} + \|\mathbf{x}_2 - \varphi^{-1}(K \varphi(\mathbf{x}_1))\|_{\infty} \quad (12)$$

où  $T$  est le nombre d'étapes temporelles dans chaque trajectoire. Les poids  $\alpha_i$  sont des hyperparamètres. L'entier  $S_p$  est un hyperparamètre pour le nombre d'étapes à vérifier dans la perte de prédiction.

Les auteurs ont introduit trois composantes d'erreur quadratique moyenne pondérée :

- précision de la reconstruction
- prédiction de l'état futur
- linéarité de la dynamique

De plus, pour tenir compte de l'absence de valeurs aberrantes dans les données, ils incorporent un terme pour pénaliser le point de données avec la perte la plus élevée. Enfin, ils ont introduit une régularisation sur les poids  $W$  pour éviter le surajustement.

## 3 Résultats et applications

### 3.1 Résultats théoriques

Dans un premier temps, nos auteurs se sont intéressés à des résultats théoriques qu'on pouvait déduire des études de Koopman et Deep Learning.

#### 3.1.1 Systèmes dynamiques pilotés par les données

On a constaté avec ces études le point fort de la dynamique basés sur les données tel que :

$$x_{k+1} = f(x_k) \tag{13}$$

Malheureusement,  $f$ , qui représente la dynamique projetant  $x$  vers l'avant dans le temps est inconnue, mais des mesures dynamiques sont disponibles.

D'un point de vue géométrique, comme dans la tradition de Poincaré, nous avons pu tirer une perspective vis à vis de la trajectoire, le point fixe et les orbites. La reformulation de  $F$  en un système d'équations différentielles donne des représentations compactes et efficaces, mais la solution reste souvent compliquée ou irreprésentable, sauf pour des classes spéciales de systèmes. L'une de ces classes est la dynamique linéaire, par exemple, un système avec  $F$  comme matrice admet une solution universelle en termes de valeurs propres grâce à l'expansion spectrale.

#### 3.1.2 Théorie des opérateurs de Koopman

En 1931 l'opérateur de Koopman  $\mathcal{K}$  a été trouvé. L'analyse de Koopman a récemment suscité une attention considérable, notamment grâce à la disponibilité croissante de données de mesure et à l'absence d'équations connues pour de nombreux systèmes.

Cette expression (3) a permis de représenter un système non-linéaire dans un cadre linéaire avec  $\mathcal{K}$  afin de pouvoir accomplir une prédiction, une estimation et un contrôle non linéaires avancé. Les représentations de dimension finie de l'opérateur de Koopman sont souvent approchées à l'aide de la DMD (Dynamic Mode Decomposition). Par construction,

la DMD identifie les structures cohérentes spatio-temporelles d'un système dynamique de haute dimension, mais elle ne capture généralement pas les transitoires non linéaires, car elle est basée sur des mesures linéaires du système  $g(x) = x$ . La eDMD et l'approche variationnelle associée de la dynamique de conformation enrichissent le modèle avec des mesures non linéaires. Il a été démontré que l'eDMD est équivalent à l'approche variationnelle de la dynamique de conformation.

L'identification de modèles de régression basés sur des mesures non linéaires entraînera généralement des problèmes, car il n'y a aucune garantie que ces mesures forment un sous-espace invariant. Une approche alternative est d'identifier les fonctions propres de  $\mathcal{K}$  satisfaisant :

$$\varphi(x_{k+1}) = \mathcal{K}\varphi(x_k) = \lambda\varphi(x_k) \quad (14)$$

Ici, les fonctions propres engendrent un sous-espace invariant et l'opérateur de Koopman produira une matrice lorsqu'il est restreint à ce sous-espace. Alors, nous avons ce que nous voulons un système avec une matrice.

Dans la pratique les fonctions propres de Koopman peuvent être plus difficiles à obtenir que résoudre directement le système, mais il permet d'obtenir une description linéaire compacte. Le défi d'identifier et de représenter les fonctions propres fournit une forte motivation pour l'utilisation de deep learning.

### 3.1.3 Koopman pour les systèmes à spectre continu

En 1932, suite à l'expansion de Koopman et Neumann, l'importance des systèmes à spectre continu a motivé plusieurs chercheurs. La pendule simple est l'un de ces cas avec une fréquence observée continue différemment des cas discrets où on a une fréquence isolée et fixe. Une avancée algorithmique a permis l'approximation des ces systèmes avec une description simple de Koopman.

Une grande innovation dans cette étude a donné deux grandes étapes : détermination de la dépendance de  $\mathcal{K}(\lambda)$ , où  $\lambda$  varie continûment (construction d'un réseau auxiliaire) et construction d'un modèle interprétable. Cela a permis d'avoir un oscillateur non-linéaire avec un spectre continu peut maintenant être représenté par une seule paire de fonctions propres conjuguées avec trajectoire en  $\cos/\sin$  parfait.

### 3.1.4 Identification les fonctions propres de Koopman

La puissance de Deep Learning pour découvrir et représenter les fonctions propres de Koopman se traduit par une représentation efficace sans surajustement et avec description minimale des dynamiques. C'est là que l'architecture conçue pour traiter une classe avec des spectres continus générés par des systèmes non linéaires comme on a expliqué précédemment entre en jeu. L'auteur entraîne le réseau par génération des trajectoires à partir de conditions



initiales aléatoires. Les modèles sont entraînés sur l'ensemble d'entraînement et comparés sur l'ensemble de validation (utilisation de 'early stopping' pour éviter le surajustement).

## 3.2 Résultats appliqués à des exemples

Dans un second temps, nos auteurs démontrent leurs approches sur plusieurs exemples dont les équations sont les suivantes.

### 3.2.1 Modèle simple avec spectre discret

$$\begin{aligned}\dot{x}_1 &= \mu x_1 \\ \dot{x}_2 &= \lambda(x_2 - x_1^2)\end{aligned}\tag{15}$$

### 3.2.2 Pendule non linéaire avec spectre continu

$$\ddot{x} = -\sin(x) \quad \Rightarrow \quad \begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = -\sin(x_1) \end{cases}\tag{16}$$

### 3.2.3 Écoulement de fluide non linéaire à haute dimension

$$\begin{aligned}\dot{x}_1 &= \mu x_1 - \omega x_2 + A x_1 x_3 \\ \dot{x}_2 &= \omega x_1 + \mu x_2 + A x_2 x_3 \\ \dot{x}_3 &= -\lambda(x_3 - x_1^2 - x_2^2)\end{aligned}\tag{17}$$

### 3.2.4 Comparaison des exemples

L'architecture correspondant à chaque exemple est donnée dans le tableau suivant.

	<b>Spectre Discret</b>	<b>Pendule avec Spectre continu</b>	<b>Écoulement de fluide 1</b>	<b>Écoulement de fluide 2</b>
<b># Couches Cachées (CC)</b>	2	2	1	1
<b>Largeur CC</b>	30	80	105	130
<b>#CC du réseau auxiliaire</b>	3	1	1	2
<b>Largeur CC du réseau auxiliaire</b>	10	170	300	20

FIGURE 3 – Architecture du réseau

Pour chaque système dynamique, nos auteurs forment plusieurs modèles dans une recherche aléatoire de l'espace des hyperparamètres et choisissent celui ayant la plus faible erreur de validation. Les meilleurs hyperparamètres sont les suivant :

	<b>Spectre Discret</b>	<b>Pendule avec Spectre continu</b>	<b>Écoulement de fluide 1</b>	<b>Écoulement de fluide 2</b>
$\alpha_1$	0.1	0.001	0.1	0.1
$\alpha_2$	$10^{-7}$	$10^{-9}$	$10^{-7}$	$10^{-9}$
$\alpha_3$	$10^{-15}$	$10^{-14}$	$10^{-13}$	$10^{-13}$
$S_p$	30	30	30	30

FIGURE 4 – Meilleures hyperparamètres

Les erreurs obtenues sont les suivantes :

	<b>Spectre Discret</b>	<b>Pendule avec Spectre continu</b>	<b>Écoulement de fluide 1</b>	<b>Écoulement de fluide 2</b>
<b>Apprentissage</b>	$1.4 \times 10^{-7}$	$8.5 \times 10^{-8}$	$5.4 \times 10^{-7}$	$2.8 \times 10^{-6}$
<b>Validation</b>	$1.4 \times 10^{-7}$	$9.4 \times 10^{-8}$	$5.4 \times 10^{-7}$	$2.9 \times 10^{-6}$
<b>Test</b>	$1.5 \times 10^{-7}$	$1.1 \times 10^{-7}$	$5.5 \times 10^{-7}$	$2.9 \times 10^{-6}$

FIGURE 5 – Tableau des erreurs

Visiblement, on a des faibles erreurs. Cependant, on voit bien que le modèle est plus adapté avec la pendule à spectre continu. Avec un certain choix de paramètres (dans le système de l'écoulement de fluide), l'ordre de grandeur de l'erreur de l'écoulement de fluide peut être pareil à celui du spectre discret et peut augmenter selon d'autre choix.

### 3.3 Applications

Dans cette section, nous allons présenter les applications dans des cas réels.

Premièrement, l'utilisation de la DMD pour la prédiction de séries temporelles. Par rapport à d'autres méthodes comme la LSTM et Hankel Linéaire, la LKIS-DMD a la plus faible erreur (RMSE) à partir de 30 étapes. La figure suivante nous montre que la valeur prédite par cette LKIS-DMD est vraiment proche de la vraie valeur à prédire.

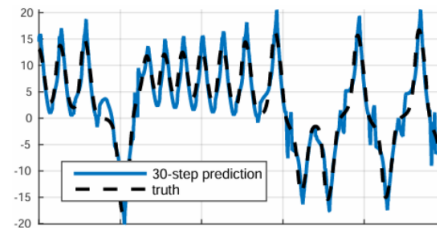


FIGURE 6 – Prédiction de séries temporelles chaotiques par la DMD

Après, on peut aussi l'utiliser dans la détection des phénomènes instables. Pour cela, l'auteur a observé les fonctions propres correspondant à la plus petite valeur propre d'une série temporelle générée par un laser infrarouge car des faibles valeurs propres correspondent à une instabilité (mode de décroissance rapide). Le modèle consiste alors à détecter les décroissances soudaines des amplitudes. Par rapport à d'autres méthodes comme l'OC-SVM et RuLSIF, le LKIS avec 10 étapes reste la meilleure en terme de score AUC (0,924 contre 0,799 et 0,803).

En tout, la DMD obtient les meilleures performances dans les études de séries temporelles liées à la dynamique (prédiction, étude des caractéristiques globales,...).

## Références

- [1] Bethany LUSCH, J. Nathan KUTZ et Steven L. BRUNTON. “Deep learning for universal linear embeddings of nonlinear dynamics”. In : *Nature Communications* 9.1 (nov. 2018), p. 4950. ISSN : 2041-1723. DOI : 10.1038/s41467-018-07210-0. URL : <https://doi.org/10.1038/s41467-018-07210-0>.
- [2] Naoya TAKEISHI, Yoshinobu KAWAHARA et Takehisa YAIRI. *Learning Koopman Invariant Subspaces for Dynamic Mode Decomposition*. 2018. arXiv : 1710.04340 [cs.LG].
- [3] B. O. KOOPMAN. “Hamiltonian Systems and Transformation in Hilbert Space”. In : *Proceedings of the National Academy of Sciences* 17.5 (1931), p. 315-318. DOI : 10.1073/pnas.17.5.315. eprint : <https://www.pnas.org/doi/pdf/10.1073/pnas.17.5.315>. URL : <https://www.pnas.org/doi/abs/10.1073/pnas.17.5.315>.